



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/731,164	12/10/2003	Naoki Kayahara	117915	5098
25944	7590	09/27/2006	EXAMINER	
OLIFF & BERRIDGE, PLC P.O. BOX 19928 ALEXANDRIA, VA 22320			LIE, ANGELA M	
			ART UNIT	PAPER NUMBER
			2163	

DATE MAILED: 09/27/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 10/731,164	Applicant(s) KAYAHARA ET AL.	
	Examiner Angela M. Lie	Art Unit 2163	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 24 July 2006.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-14 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-14 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 10 December 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Claim Rejections - 35 USC § 112

1. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.
2. Claims 1, 6, 9, 12 and 13 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.
3. Furthermore, in the above listed claims the applicant states that "a similarity computing device to acquire a plurality of documents", but it is not specified based on what criteria those documents are selected. The applicant needs to clearly define what or why the specific documents are selected.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. **Claims 1, 6, 9 and 12-14 are rejected under 35 U.S.C. 102(b) as being anticipated by Seki et al (US Publication 20020143737).** Seki discloses a document extracting device and a method, comprising: a similarity computing device to acquire

plurality of documents to be candidates for extraction (Figure 9, elements 903 and 904) and computing all degrees of similarity between the documents (paragraph 47); and a document extracting device to extract a combination of documents from the candidate documents with a sum of the degrees of similarity between the documents computed by the similarity computing device is the smallest when any number of documents are extracted from among a group of the documents (paragraph 49, so that if the documents are similar (duplicate), only one document will be outputted to the user, but if those document have small similarity then both of them will be outputted).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. **Claims 2-5, 7, 8 and 10-11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Seki et al (US Publication 2002/0143737) in the view of Wyard et al (US Patent 6167398).**

As to claims 2, 3, 7 and 10, Seki teaches the similarity computing device and the method comprising: a mutual similarity computing functional unit to compute the similarity between the documents, however he does not teach that the search is based on the document vector involving a character-string frequency computing and he does not explicitly state that the document is divided into predetermined character strings via

any of the: a morphological analysis method, an n-gram method and a stop-word method. Wyard teaches an information retrieval system and method that generates weighted comparison results to analyze the degree of dissimilarity between documents, wherein the word frequency corresponds to character-string frequency, the document is divided into equal portions (column 9, lines 53-59) and an n-gram method is used (paragraph 3, lines 36-41). It would have been obvious to one of the ordinary skill in the art during the time the invention was made to use inverse document frequency technique (i.e. document vector) to measure the similarity as taught by Wyard (column 2, lines 5-26) and to divide the document into sections and to use a well known n-gram method (Seki also acknowledges that the certain portions of the document should be treated separate, for instance title (paragraph 64)) in Seki's similarity determining means because inverse document frequency method/algorithm is well know in the art and it is relying on term weighting that improves the performance and furthermore it is also easy to compute and therefore it is faster than Boolean search or comparison, further certain portions of the document might be weighted differently to find a better match, for instance the title of the document might be given highest weight (most significant), and therefore has to be identified or separated from the rest of the passage.

As to claim 4, Wyard further teaches the character –string frequency computing functional unit generating document vector obtained by weighting each of the documents by TFIDF on the basis of the frequency of appearance of the divided character strings (column 8 lines 22-29 and lines 47-55).

As to claim 5, Wyard further teaches the mutual similarity computing functional unit computing the degrees of similarity between the documents by a vector space method on the basis of the document vectors of the documents (column 8, lines 22-29).

As to claim 8, Wyard further teaches the similarity computing device comprising: a character-string-dividing function to divide each of the documents into character strings using any one of character string division methods (column 3, lines 36-41); a character-string frequency computing function to generate document vectors obtained by weighting each of the documents by TFIDF (column 8, lines 47-55) on the basis of the frequency of appearance of the divided character strings; and a mutual similarity computing function to compute the degrees of similarity between the documents by a vector space method on the basis of the document vectors of the documents (column 8, lines 22-29).

As to claim 11, Seki does not teach that the search is based on the document vector involving a character-string frequency computing and in particular TFIDF. Wyard teaches an information retrieval system and method that generates weighted comparison results to analyze the degree of dissimilarity between documents, wherein the word frequency corresponds to character-string frequency. Wyard also teaches using n-gram string division method (column 3, lines 36-41) and weighting each of the documents by TFIDF on the basis of the frequency of occurrence particular character string. It would have been obvious to one of the ordinary skill in the art during the time the invention was made to use inverse document frequency technique (i.e. document vector) to measure the similarity as taught by Wyard (column 2, lines 5-26) in Seki's

similarity determining means because inverse document frequency method/algorithm is well known in the art and it is relying on term weighting that improves the performance and furthermore it is also easy to compute and therefore it is faster than Boolean search or comparison.

Response to Arguments

8. Applicant's arguments filed July 24, 2006 have been fully considered but they are not persuasive.

9. The claim rejection under 35 U.S.C 101 regarding claims 6-8 and 13, has been withdrawn in response to the amendment made in the independent claim 6.

10. The objection regarding claims 4, 8, and 11, also has been withdrawn.

11. With regards to the rejection under 35 U.S.C 112 second paragraph, the applicant did not amend the claims 1,6,9, 12 and 13, in such a manner that would overcome 35 U.S.C 112 second paragraph. On page 7, last paragraph applicant states that the corresponding claims were amended to specify: acquiring "a plurality of documents from a information source according to a predetermined criteria", however this limitation is not incorporated in any of the independent claims.

12. With respect to the applicant's assertion on page 8, forth paragraph stating that Seki does not disclose the use of a sum of degrees of similarities between the candidate documents, as required by the independent claims. The examiner disagrees with this assertion, as stated in paragraph 47, the content of summaries of the documents (candidates) is compared, and then measured (for instance 80% similarity,

which indeed is sum of degrees, because as shown in figure 8, there are multiple rules for establishing if there are duplicate documents, therefore all those rules constitute sum of degrees i.e. matching full length of the text (paragraph 46), comparing summaries of the documents (paragraph 47)).

13. Furthermore, also on page 8, fifth paragraph, the applicant alleges that Seki does not make it possible to eliminate overlapping when the number of communication information is limited. The examiner would like to note few points; most importantly the limitation argued by the applicant is not in any of the claims, further the applicant's assertion is not supported, because certainly Seki did not state that limited information would prevent eliminating duplicates.

The Prior Art

14. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Gomes et al (US Patent 6615209) disclose a method for comparing two documents for the similarity.

Conclusion

15. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

16. A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Inquiry

17. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Lie whose telephone number is 571-272-8445. The examiner can normally be reached on M-F.

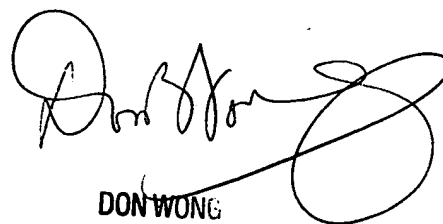
18. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Don Wong can be reached on 571-272-1834. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2163

19. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.



Angela M Lie



DON WONG
SENIOR ADVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100